

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 05-128157

(43)Date of publication of application : 25.05.1993

(51)Int.Cl.

G06F 15/40

G06F 12/00

(21)Application number : 03-292660

(71)Applicant : RICOH CO LTD

(22)Date of filing : 08.11.1991

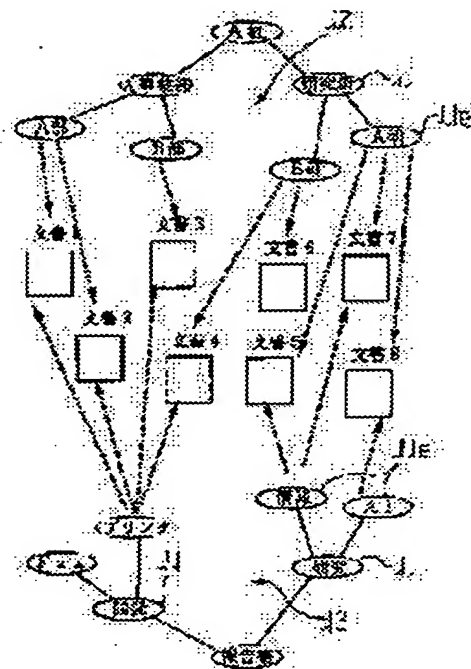
(72)Inventor : IWASAKI MASAJIROU

(54) DOCUMENT RETRIEVAL SYSTEM

(57)Abstract:

PURPOSE: To correct lack of collective retrieval function of the data that satisfies a certain condition, which is a defect of a hyper text in which plural document data are stored and, by specifying a key word in a displayed document, related documents are retrieved and then retrieving result is displayed.

CONSTITUTION: A classification tree data storing means in which the data of plural classification trees 12 and 13 in which respective nodes 11 linked to corresponding documents are connected with each other in tree-like manner for document classification and management are stored is provided. in addition, a retrieval processing means which, with a specified arbitrary node 11 as a starting point for retrieval, retrieves the document belonging to the three below that node for the specified classification trees 12 or 13 is provided. So, at retrieval using a hyper text function, collective retrieval, within proper range, can be made with ease with the data satisfying certain conditions.



LEGAL STATUS

[Date of request for examination] 13.11.1996

[Date of sending the examiner's decision of rejection] 10.11.1998

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's
decision of rejection] 10-19481

[Date of requesting appeal against
examiner's decision of rejection] 10.12.1998

[Date of extinction of right]

Copyright (C); 1998,2003 Japan Patent Office

(19)日本国特許庁(JP)

(12) 公開特許公報(A)

(11)特許出願公開番号

特開平5-128157

(43)公開日 平成5年(1993)5月25日

(51)Int.Cl.⁵

G 0 6 F 15/40
12/00

識別記号

5 0 0 L 7060-5L
5 4 7 H 7832-5B

庁内整理番号

F I

技術表示箇所

審査請求 未請求 請求項の数7(全14頁)

(21)出願番号 特願平3-292660

(22)出願日 平成3年(1991)11月8日

(71)出願人 00006747

株式会社リコー

東京都大田区中馬込1丁目3番6号

(72)発明者 岩崎 雅二郎

東京都大田区中馬込1丁目3番6号 株式
会社リコー内

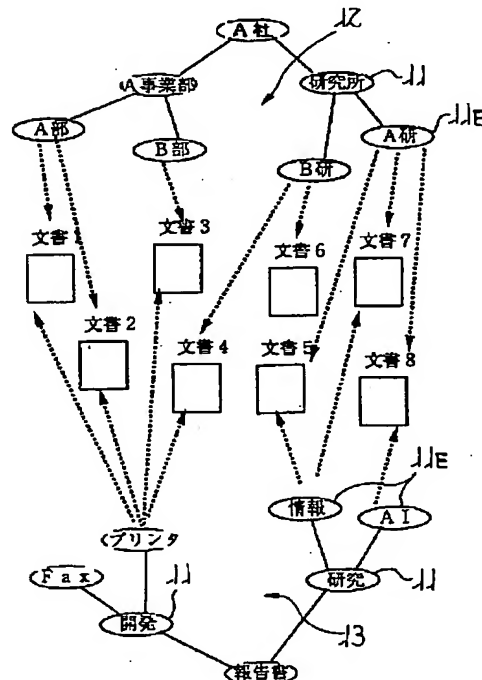
(74)代理人 弁理士 柏木 明

(54)【発明の名称】 文書検索装置

(57)【要約】

【目的】 複数の文書データを記憶し、表示されている文書中のキーワードを指定することにより関連文書を検索し検索結果を表示させるようにしたハイパーテキストの欠点である、ある条件を満たすデータの一括検索機能の欠如を改良する。

【構成】 対応する文書へリンクさせた各ノード11を樹状に接続して文書を分類管理するための複数の分類ツリー12、13のデータを格納した分類ツリーデータ記憶手段を設け、選択指定された分類ツリー12又は13について指定された任意のノード11を検索起点としてそのノード下位のツリーに属する文書を検索する検索処理手段を設け、ハイパーテキスト機能による検索時に、ある条件を満たすデータの適切な範囲の一括検索を容易に行なう機能を持たせた。



【特許請求の範囲】

【請求項1】 複数の文書データを記憶し、表示されている文書中のキーワードを指定することにより関連文書を検索し検索結果を表示させるようにした文書検索装置において、対応する文書へリンクさせた各ノードを樹状に接続して文書を分類するための複数の分類ツリーデータを格納した分類ツリーデータ記憶手段を設け、選択指定された分類ツリーについて指定されたノードを検索起点としてそのノード下位のツリーに属する文書を検索する検索処理手段を設けたことを特徴とする文書検索装置。

【請求項2】 分類ツリーデータ記憶手段を、各分類ツリー、各分類ツリーの各ノードにリンクされた文書の題名及びその文書の内容要約のデータを格納したものとし、検索処理手段による検索時にこの分類ツリーデータ記憶手段に記憶された内容の一覧を選択自在に表示させる対話的処理手段を設けたことを特徴とする請求項1記載の文書検索装置。

【請求項3】 複数の文書データを記憶し、表示されている文書中のキーワードを指定することにより関連文書を検索し検索結果を表示させるようにした文書検索装置において、文書データと独立して設定されてキーワードと文書とを関連付けたリンクデータを格納したリンクデータ記憶手段を設け、指定されたキーワードについて前記リンクデータに基づき文書を検索する検索処理手段を設けたことを特徴とする文書検索装置。

【請求項4】 リンクデータ中に、キーワードと文書との関連度情報を伴いユーザ毎に設定されて選択自在なリンクデータを含むことを特徴とする請求項3記載の文書検索装置。

【請求項5】 各々別個のリンクデータを持つ複数種のリンク集合を階層化させて選択指定自在にリンクデータ記憶手段に格納したことを特徴とする請求項3記載の文書検索装置。

【請求項6】 リンクデータとともにリンクデータ記憶手段に格納されたキーワードと文書との関連度情報を検索結果に応じたユーザ指示により修正する関連度情報学習手段を設けたことを特徴とする請求項3記載の文書検索装置。

【請求項7】 複数の文書データを記憶し、表示されている文書中のキーワードを指定することにより関連文書を検索し検索結果を表示させるようにした文書検索装置において、対応する文書へリンクさせた各ノードを樹状に接続して文書を分類するための複数の統制分類ツリーデータを格納した分類ツリーデータ記憶手段を設け、キーワードとフリー分類ツリーのノードとを関連付ける分類テーブルを設け、文書データを形態素解析して自動的にキーワードを抽出するキーワード抽出部を有してこの形態素解析により得られたキーワードによるフリー分類ツリーに文書を登録する文書登録手段を設け、選択指定

された分類ツリーについて指定されたノードを検索起点としてそのノード下位のツリーに属する文書を検索する検索処理手段を設けたことを特徴とする文書検索装置。

【発明の詳細な説明】

【0001】

【産業上の利用分野】本発明は、文書検索システム、特にハイパーテキストシステムにおける文書検索装置に関する。

【0002】

【従来の技術】近年、文書などのテキスト情報の典型的な利用形態として注目を集めているものに、「ハイパーテキスト」と称されるものがある。これは、論文「ハイパーメディアとは何か」(Computer Today, 1990/7 No. 38)中に記載されているように、複数のテキスト情報を種々の参照関係を表すリンクによって非線形な形に関連付けたものである。即ち、文書中の単語(キーワード)からその単語に関連する文書へリンクが張られており、ユーザはその単語を指定することにより関連文書を参照することができる。同様に、その文書中の単語からも他の文書へリンクが張られており、このようなリンクを順次辿ることにより、ユーザは目的とする文書を得ることができる。

【0003】

【発明が解決しようとする課題】ところが、このようなハイパーテキストは、上記論文中にも記載されているように、データベースシステムのようにある条件を満たすデータを一括して検索するという機能が弱く、リンクを各々の文書中の単語毎に設定するために膨大なリンクを設定しなければならない。

【0004】また、ハイパーテキストは関連のある文書を順に辿っていくことにより必要とする文書を検索するため、辿っていく過程が複雑な場合には自分がどこにいるか判らなくなってしまう、不慣れな人では途惑い、混乱を生じて、検索しにくいものである。また、ユーザの興味や仕事の内容などによっては、ユーザ毎或いはグループ毎に文書間のリンクが異なっている必要があるが、それを統一的に管理することができず、適切な検索ができない。さらには、文書間の関連度もユーザ毎或いはグループ毎に異なり、かつ、各関連度が変化する可能性があるが、従来のハイパーテキストではこれに対処できず、やはり、適切な検索ができない。よって、文書が多い場合には検索時間がかかってしまうものとなる。

【0005】

【課題を解決するための手段】請求項1記載の発明では、複数の文書データを記憶し、表示されている文書中のキーワードを指定することにより関連文書を検索し検索結果を表示させるようにした文書検索装置において、対応する文書へリンクさせた各ノードを樹状に接続して文書を分類するための複数の分類ツリーデータを格納した分類ツリーデータ記憶手段を設け、選択指定された分

類ツリーについて指定されたノードを検索起点としてそのノード下位のツリーに属する文書を検索する検索処理手段を設けた。

【0006】この際、請求項2記載の発明では、分類ツリーデータ記憶手段を、各分類ツリー、各分類ツリーの各ノードにリンクされた文書の題名及びその文書の内容要約のデータを格納したものとし、検索処理手段による検索時にこの分類ツリーデータ記憶手段に記憶された内容の一覧を選択自在に表示させる対話的処理手段を設けた。

【0007】また、請求項3記載の発明では、複数の文書データを記憶し、表示されている文書中のキーワードを指定することにより関連文書を検索し検索結果を表示させるようにした文書検索装置において、文書データと独立して設定されてキーワードと文書とを関連付けたリンクデータを格納したリンクデータ記憶手段を設け、指定されたキーワードについて前記リンクデータに基づき文書を検索する検索処理手段を設けた。

【0008】この際、請求項4記載の発明によれば、リンクデータ中に、キーワードと文書との関連度情報を伴いユーザ毎に設定されて選択自在なリンクデータを含むものとした。

【0009】また、請求項5記載の発明では、各々別個のリンクデータを持つ複数種のリンク集合を階層化させて選択指定自在にリンクデータ記憶手段に格納した。

【0010】さらに、請求項6記載の発明では、リンクデータとともにリンクデータ記憶手段に格納されたキーワードと文書との関連度情報を検索結果に応じたユーザ指示により修正する関連度情報学習手段を設けた。

【0011】一方、請求項7記載の発明では、複数の文書データを記憶し、表示されている文書中のキーワードを指定することにより関連文書を検索し検索結果を表示させるようにした文書検索装置において、対応する文書へリンクさせた各ノードを樹状に接続して文書を分類するための複数の統制分類ツリーデータを格納した分類ツリーデータ記憶手段を設け、キーワードとフリー分類ツリーのノードとを関連付ける分類テーブルを設け、文書データを形態素解析して自動的にキーワードを抽出するキーワード抽出部を有してこの形態素解析により得られたキーワードによるフリー分類ツリーに文書を登録する文書登録手段を設け、選択指定された分類ツリーについて指定されたノードを検索起点としてそのノード下位のツリーに属する文書を検索する検索処理手段を設けた。

【0012】

【作用】請求項1記載の発明によれば、文書を管理する複数の分類ツリーにより検索可能としたので、その検索範囲の設定・変更が、分類ツリーの選択指定及び検索起点となるノードの指定により自在となり、ハイパーテキストシステムにおいて、ある条件を満たすデータの適切な範囲の一括検索を容易に行なう機能を持たせることが

できる。この時、複数の分類ツリーの指定利用により、複合条件についての検索も可能となる。この際、請求項2記載の発明によれば、分類ツリーやユーザの指定したノードに含まれる文書データの書誌的事項等の一覧が表示されて対話処理可能とされているので、検索個所がわかりやすく、不慣れな人でも検索しやすいものとなる。

【0013】同様に、請求項3記載の発明による場合も、キーワードと文書とを関連付けたリンクデータを文書データとは別個に用意し、キーワードによる一括検索を可能としたので、ハイパーテキストシステムにおいて、ある条件を満たすデータの一括検索を容易に行なう機能を持たせることができる。この時、リンクデータは文書データとは独立したものであり、各文書に付けられているキーワードから全文書共通として自動的に生成し得るので、文書数が多い場合であってもリンクの生成が容易なものでもある。この際、請求項4記載の発明によれば、このようなリンクデータ中にユーザ毎に設定したものも選択自在に含まれるので、ユーザ毎に適切な検索が可能となる。さらに、請求項5記載の発明によれば、リンクデータを複数種のリンク集合として持たせたので、ユーザ毎、グループ毎といった単位での適切な検索が可能となる。特に、このようなリンク集合が階層化されて格納されているので、リンク集合の扱いが各層単位で統一的となり容易になる。また、請求項6記載の発明によれば、検索結果に応じたユーザ指示によりリンクデータ中の関連度情報を学習修正自在としたので、ユーザやグループなどの仕事や興味の変化などに対応した適切な検索が可能となる。

【0014】また、請求項1記載の発明による場合も請求項3記載の発明による場合も、分類ツリーデータ、リンクデータが何れも文書に対するリンクを持っているので、文書検索を短時間で済ませる。

【0015】さらに、請求項7記載の発明によれば、ユーザ登録による統制分類ツリーとは別個に、文書登録時にキーワード抽出を利用して文書を自動的にフリー分類ツリーに登録分類して、統制分類ツリーと併用自在としたので、文書の分類形態が細かなものとなり、これらの分類ツリーの併用により検索時の絞り込みが容易で検索処理も短時間で済むものとなる。

【0016】

【実施例】本発明の第一の実施例を図1ないし図9に基づいて説明する。まず、図2は本実施例の基本構成を示すブロック図であり、大別すると、入力部1と記憶部2と処理部3と出力部4とにより構成されている。入力部1はキーボードによりキーワード等を入力し、マウス等のポインティングデバイスによりキーワード、ノード等の選択を行なうものである。記憶部2は、複数の文書データを記憶した文書データ記憶部5の他、後述する分類ツリーデータを記憶した分類ツリーデータ記憶部(分類ツリーデータ記憶手段)6と、リンクデータを記憶した

リンクデータ記憶部（リンクデータ記憶手段）7とを有する。処理部3中には検索処理部（検索処理手段）8と文書表示部9とを有する。この内、検索処理部8は入力部1から得られた情報と記憶部2中の分類ツリーデータとリンクデータとを用いて、文書データ中から文書を検索する。また、文書表示部9は検索した文書を表示し、前述したようにハイパーテキストとしての表示機能を持つものである。出力部4はディスプレイに検索結果等を出力する。

【0017】文書データ記憶部5に記憶される文書データ10の構造は、図3に示すように、文書データとともに、文書番号、N個（Nは任意数）のキーワードを持つ。

【0018】また、本実施例では、図1に示すように複数の文書（図示例は、文書1～文書8）を分類管理するための、幾つかの分類ツリーが設定されている。例えば、報告書の場合であれば、報告書の表紙に記載された項目毎に分類ツリーが用意される。つまり、報告者の所属による分類ツリー、報告内容の分類ツリー、報告書番号による分類ツリー、報告書作成日による分類ツリー、プロジェクトによる分類ツリーなどが考えられる。これらの分類ツリーはノード11を樹状に接続したものであり、分類ツリーの各末端ノード11eは文書エントリを持ち、対応する文書へリンクしている。よって、各文書は各分類ツリーによって属するノードが変わることになる。

【0019】例えば、図1中の分類ツリー12は会社の構成を表し、各ノード11は対応する部署に属する文書にリンクしている。また、分類ツリー13は報告書の内容を表し、各ノード11は対応する項目に属する文書にリンクしている。よって、図示例であれば、文書8は会社構成を表す分類ツリー12中の部署「A研」で示すノード11eに属するとともに、報告書内容を表す分類ツリー13中の項目「A1」で示すノード11eに属していることが判る。

【0020】このような分類ツリーに関するデータは、前記分類ツリーデータ記憶部6中に格納されている。分類ツリーの各ノード11は、自ノードの分類項目名（ノード名）及び下位ノードへのリンク情報を持つ。また、末端ノード11eでは自ノードの分類項目名及びその分類項目に属する文書テーブルへのポインタを持つ。文書テーブルの文書エントリ14は、図4に示すように、文書データのID、その文書のタイトル及び内容要約文の情報を持つ。

【0021】このような分類ツリーのデータは、各文書登録時に生成される。即ち、各文書には分類ツリー毎のノード名（例えば、報告書の所属、報告内容、報告書番号、プロジェクト名など）が人為的に付与されており、本実施例の検索装置が対応する分類ツリーの末端ノード11eに文書エントリ14を追加する。

【0022】このような分類ツリーを利用することにより、検索範囲を設定できる。検索起点として、あるノードを指定すると、検索はそのノードから下位のツリー（サブツリー）に属する文書に対して行なわれる。このような検索範囲は、検索起点のノードを変更指定することにより、容易に変更できる。これは、従来のリレーショナルデータベースのように検索範囲をコマンドラインから入力するのに比して、検索範囲の指定・変更の操作が容易なものである。また、複数の分類ツリーを併用し、各々の検索起点となるノードを指定すれば、各々のサブツリーに属する文書をAND又はOR条件なる複合条件で検索できるものとなる。例えば、図1に示した例で、「A事業部」に関連する「開発」関連の文書を検索したい場合には、分類ツリー12中の部署「A事業部」なるノード11からのサブツリーに属し、かつ（又は）、分類ツリー13中の項目「開発」なるノード11からのサブツリーに属する文書を検索すればよいものとなる。

【0023】ここに、分類ツリーを利用した検索画面（対話的処理手段）15の例を図5に示す。画面15の上部15aには、分類ツリーが表示され、下部15bには文書の内容一覧が表示され、対話的な選択操作に基づき検索が可能とされている。図中、分類ツリー中で「研究所」と「研究」なるノードが四角の枠で囲まれているが、これは、分類ツリー12、13について「研究所」、「研究」なるノード11を各々検索起点として指定したことを表す。このような指定された各ノードのサブツリーに属する文書の一覧が下部15bに文書の題名（タイトル）、内容（文書の内容要約）により表示される。このような文書一覧表示の中から文書を指定することにより、文書検索が可能となる。図示例では、2つの分類ツリー12、13しか表示されていないが、さらに、複数の分類ツリーをも表示させ、各分類ツリーのノードを指定した結果、得られるサブツリーに属する文書を一覧表示させて選択に供するようにしてもよいことはもちろんである。

【0024】即ち、分類ツリーを利用した検索処理及び表示内容をまとめると、

- ① ユーザは分類ツリーを指定する。
- ② 分類ツリーデータにより分類ツリーを検索画面15に表示する。
- ③ ユーザは表示された分類ツリー中から検索起点とするノードを指定する。
- ④ 指定されたノードを検索起点として、下位のサブツリーを辿り、末端ノードから文書タイトル及び内容要約文データの一覧を検索画面15に表示する（複数の分類ツリーのノードを指定した場合には、各分類ツリーの指定されたノードのサブツリーから文書を検索し、各サブツリーで検索された文書のAND/ORをとり、検索結果として内容を検索画面15に一覧表示する）。

- ⑤ ユーザは表示された内容一覧から文書を選択する。
 ⑥ 選択された文書に対応する分類ツリーの末端ノードから文書へのリンク情報を得て、文書を取り出し、検索画面15に表示する。

【0025】このような検索に際しては、文書を分類管理する分類ツリーを表示し、かつ、ユーザの指定したノードの下位サブツリーに含まれる文書内容の一覧が表示されるので、検索している個所が判りやすく、不慣れな人でも混乱することなく容易に検索し得るものとなる。

【0026】ところで、本実施例では、図6に示すような構造のリンクデータ16がリンクデータ記憶部7に格納されている。各キーワードはそれをキーワードとする文書とリンクしており、リンクデータ16は各キーワード毎にリンクしている文書の文書番号、文書名及びキーワードと文書との関連度情報とからなる。このようなリンクデータ16は全文書に付けられたキーワードの情報をキーワード毎に集めることにより自動的に生成される。

【0027】このようなリンクデータ16を用いたリンクの様子を図7に示す。従来のハイパーテキストでは文書毎に単語から他の文書へのリンクデータを持っていたが、本実施例では、図示例からも判るように、リンクデータ16は文書データとは独立したものとし、全文書に共通とすることにより、簡単にハイパーテキストのリンクを生成することが可能となる。このようなリンクデータ16は複数種持つことができる。例えば、図示例のように、全てのユーザが共通に利用できるパブリックリンクデータと、各ユーザが個別に利用するプライベートリンクデータというように設定できる。このようなリンクデータ設定によれば、全てのユーザが一般的に持っているキーワードと文書の関連（パブリックリンク）だけでなく、プライベートリンクにより各ユーザの興味に対応したキーワードのリンクを構成できる。図7では、文書AはキーワードAを含んでいるが、キーワードAはパブリックリンクデータ16_nにより文書1～文書3にリンクし、プライベートリンクデータ16_nにより文書4、文書5にリンクしている。

【0028】このようなプライベートリンクデータ16_nをユーザ毎に構築し指定し得るので、ユーザの興味などに応じた適切な検索が可能となる。

【0029】このような分類ツリー、リンクデータを利用して、検索処理部8では図8に示すフローチャートに従い下記のように検索処理を行なう。

- ① ユーザは、キーワード、検索に利用する分類ツリー、検索起点とするノード及び検索条件（AND又はOR）を入力部1より指定する（指定しない時は、前回指定された内容が用いられる）。
 ② 与えられたキーワードの各リンクデータ16（より具体的には、パブリックリンクデータ、プライベートリンクデータ等）を検索する。検索した文書集合を検索集

合Aとする。

③ 指定された分類ツリーの検索起点となるノードからの下位サブツリーに対して文書を検索する。検索した文書集合を検索集合Bとする。

④ 指定された検索条件に従い、検索集合Aと検索集合BとのAND又はORをとり、結果を検索結果とする。

⑤ 検索結果の文書についてリンクデータ中の関連度の大きい順に文書名を表示する。ただし、検索結果の文書が一つの場合には、この処理を行わず、⑦の処理に進む。

⑥ 検索結果から、ユーザが参照したい文書名を選択する。

⑦ 選択された文書について、その文書番号により文書データ記憶部5中を検索してその文書データを表示する。この処理は、文書表示部9が行なう。

【0030】文書表示部9では、検索した文書を表示し、ハイパーテキストの機能として、表示中の文書中のキーワードを表示し、ユーザがその中からキーワードを指定する。そして、前述したように検索処理部8によりキーワードから関連する文書を検索する。

【0031】このような文書表示処理をまとめると、

① 検索処理部8で指定された文書の文書番号から文書データを検索する。

② 検索された文書データ中の文書を表示する。

③ 予め決められているキーボードのキーを押すことにより、表示されている文書中の全てのキーワードを反転表示等により明示する。ユーザはこのようなキーワードの中から所望のものを選択する。

【0032】ところで、キーワードと文書とを関連付けるリンクデータ16には図6に示したように、キーワードと文書との間の関連度情報も付与されているので、ユーザは、この関連度により検索された文書の関連の度合いを知ることができる。文書の関連度は、例えば0～100の値で示される。そこで、あるキーワードにより検索した文書がユーザの要求していた文書と一致していた（キーワードと文書との関連度が強い）場合には、ユーザが適切な検索結果であることを本実施例の検索装置に知らせることにより、本検索装置中の関連度情報学習手段（図示せず）はそのリンクデータ16中の関連度を適

当に、例えば5ポイント程度、上げるように修正する。一方、キーワードによって検索した文書がユーザの要求していた文書と一致しない（キーワードと文書との関連度が弱い）場合には、ユーザが不適切な検索結果であることを本検索装置に知らせると、関連度情報学習手段はそのリンクデータ16中の関連度を適当に、例えば5ポイント程度、下げるように修正する。このように検索結果に応じたユーザの適否の指示により関連度情報が修正学習されるので、キーワードに対する文書の関連度の順位の変更ができ、ユーザの興味の変化などに対応したリンクデータとし、より適切な検索が可能となる。

【0033】なお、リンクデータで示されるキーワードと文書とのリンクが全く不適当と思われる場合には、リンクデータからその文書エントリを削除し、また、逆に、適当と思われる文書が欠落している場合にはそのリンクデータに文書エントリを加えるようにしてもよい。

【0034】つづいて、本発明の第二の実施例を図10及び図11により説明する。本実施例は、図10に示すようにリンクデータ16の種類をさらに多様化し、例えば全てのユーザが利用できるパブリックリンクデータ16_hと、グループ毎に利用できるグループ別リンクデータ16_gと、各ユーザ毎に利用できるユーザ別リンクデータ16_uとに区分したものである。このような各リンクデータは図11に示すようなリンクツリーとして階層化されている。このリンクツリーにおける各ノードのリンクは重複しないように設定されている。即ち、パブリックリンクデータ16_hで指定されているリンクデータはその下位層のグループ階層やユーザ階層では定義されることはない。

【0035】よって、ユーザが利用するノードをこのリンクツリー階層構造のノードから適切に選択することにより、適切な文書のリンクを構成することができる。例えば、ユーザがパブリックとユーザAとのノードを指定すれば、汎用のリンク及びグループAに関連し、かつ、ユーザAが個人的に生成したリンクにより文書を検索できる。また、グループAを指定した場合には、グループAに関連したリンクにより文書を検索できる。

【0036】よって、リンクを管理する上で、ユーザの興味や仕事の内容などによって、ユーザ毎、グループ毎でリンクが異なる必要があるが、図11に示すような階層型構造によれば、各層単位で統一的な管理が可能となる。また、本実施例について前述したようなリンクデータ中の関連度情報の学習機能を持たせれば、ユーザやグループに応じた関連度情報の確保が可能となる。

【0037】さらに、本発明の第三の実施例を図12ないし図15により説明する。まず、本実施例では、前述した実施例中で示した分類ツリーを統制分類ツリーとして扱う一方、このような統制分類ツリーとは別個にフリー分類ツリーという概念を導入したものである。即ち、図1等 に示した統制分類ツリーは文書登録時にユーザによって人為的に生成登録されたものであり、文書の基本的な分類でしかなく、文書の内容について細かく分類した分類ツリーを作成したとしても、ユーザが文書を登録する際に文書を読み、その文書がどのノードに属するかを判断する必要があり、登録作業が困難である。これに対して本実施例で導入したフリー分類ツリーは、統制分類ツリーと似てはいるものの、文書の登録時に文書データを形態素解析してキーワードを自動的に抽出し、フリー分類ツリーに自動的に文書を登録するようにしたものである。

【0038】このため、全体構成においては、図12に示すように、記憶部2中には分類ツリーデータ記憶部6にフリー分類ツリーを登録する際に利用する分類テーブル17が追加され、その登録処理を行なう文書登録部（文書登録手段）18が処理部3中に追加されている。この文書登録部18はキーワード抽出部（図示せず）を有する。

【0039】また、フリー分類ツリーにおける末端ノード19_eには、図13に示すように、自ノードの分類項目名（ノード名）、その分類項目に属する文書テーブルへのポインタ及びそのノード19_eの内容に含まれるキーワードを持つ。フリー分類ツリーにおける文書エントリ20は、図14に示すように、文書データのID、その文書のタイトル、内容要約文（即ち、図4に示した統制分類ツリーの文書エントリ14と同じ）に加え、文書のキーワードが付与されている。このキーワードは文書のキーワードであるとともに、図13に示したように、ノードの内容に含まれるキーワードである。

【0040】フリー分類ツリーでは、文書登録時に各文書を形態素解析し、単語に分割し、分割抽出された単語（キーワード）に対応する分類データを分類テーブル17から取出す。この分類テーブル17は図15に示すように各キーワードがどのフリー分類ツリーのどのノードに属するかを示すものであり、取出された分類データで示されるフリー分類ツリー及びそのノードに、抽出されたキーワードとともに文書を登録する。つまり、フリー分類ツリーについては、キーワード付けやフリー分類ツリーへの登録作業は不要である。

【0041】検索に際して、このようなフリー分類ツリーを統制分類ツリーとともに併用することにより、統制分類ツリーのみによる場合よりも絞り込みが容易となり、検索時間が短縮される。

【0042】ところで、統制分類ツリーの場合には、一つの分類ツリー中では、ある一つの文書が一つのノードにしか存在しないが、フリー分類ツリーの場合には、文書から多数のキーワードが抽出され、各々についてフリー分類ツリーに登録されるので、一つの分類ツリー中であっても、ある一つの文書が複数のノードに存在し得ることになる。即ち、一つのノードが大量の文書を持つことも考えられ、フリー分類ツリー中の検索起点となるノードの選択だけでは、文書を十分に絞り込むことができない場合もある。このような場合には、複数の統制分類ツリーと複数のフリー分類ツリーとにより検索集合を絞り込むとともに、ユーザがキーワードを指定し、フリー分類ツリー中のノードに含まれる文書エントリのキーワードと衝合することにより文書を検索すればよい。これは、前述したように、ノードに対応付けられているキーワードは、予め分類テーブル17から抽出されており、各ノードに付与されているからである。

【0043】このようなフリー分類ツリーを併用した場

合の検索処理も、基本的には、統制分類ツリーによる場合と同様に行なわれる。即ち、図5に示した検索画面15を参照すると、上部15aには統制分類ツリー及びフリー分類ツリーが表示される。そこで、フリー分類ツリーを併用した検索処理及び表示内容をまとめると、

- ① ユーザは統制分類ツリー及びフリー分類ツリーを指定する。
- ② 各分類ツリーデータにより統制分類ツリー及びフリー分類ツリーを検索画面15に表示する。
- ③ ユーザは表示された分類ツリー中から検索起点とするノードを指定する。
- ④ 統制分類ツリーの場合、指定されたノードを検索起点として、下位のサブツリーを辿り、末端ノードから文書集合Aを得る（複数の統制分類ツリーのノードを指定した場合には、各統制分類ツリーの指定されたノードのサブツリーから文書を検索し、各サブツリーで検索された文書集合のAND/ORをとり、文書集合Aとする）。
- ⑤ フリー分類ツリーの場合、指定されたノードを検索起点として、下位のサブツリーを辿り、末端ノードの文書エントリから文書集合Bを得る（複数のフリー分類ツリーのノードを指定した場合には、各フリー分類ツリーの指定されたノードのサブツリーから文書を検索し、各サブツリーで検索された文書集合のAND/ORをとり、文書集合Bとする）。
- ⑥ 文書集合Aと文書集合BとのANDをとり、得られた文書集合について、文書エントリの文書タイトル及び内容要約文を検索画面15に一覧表示する。
- ⑦ 文書集合Cが多過ぎる場合には、ユーザが適当なキーワードを与え、文書集合Cを絞り込む。即ち、
 - (a) ユーザが適当なキーワード（ユーザが指定したフリー分類ツリー中のノードの各文書エントリに付与されているキーワード）を指定する。
 - (b) ユーザが指定したキーワードと文書集合Cのフリー分類ツリーの文書エントリに付与されているキーワードが一致しない文書を文書集合Cから削除する。
 - (c) 新たにこの文書集合Cについて、文書エントリの文書タイトル及び内容要約文を一覧表示する。
- ⑧ ユーザは表示された内容一覧から文書を選択する。
- ⑨ 選択された文書に対応する分類ツリーの末端ノードから文書へのリンク情報を得て、文書を取り出し、検索画面15に表示する。

【0044】

【発明の効果】本発明は、上述したように構成したので、請求項1記載の発明によれば、文書を分類管理する複数の分類ツリーにより検索可能としたので、その検索範囲の設定・変更が、分類ツリーの選択指定及び検索起点となるノードの指定により自在となり、ハイパーテキストシステムにおいて、ある条件を満たすデータの適切な範囲の一括検索を容易に行なう機能を持たせることが

でき、この時、複数の分類ツリーの指定利用により、複合条件についての検索も可能となる。この際、請求項2記載の発明によれば、分類ツリーやユーザの指定したノードに含まれる文書データの一覧を表示させて対話処理可能としたので、検索している個所がわかりやすく、不慣れな人でも検索しやすいものとする事ができる。

【0045】同様に、請求項3記載の発明による場合も、キーワードと文書とを関連付けたリンクデータを文書データとは別個に用意し、キーワードによる一括検索を可能としたので、ハイパーテキストシステムにおいて、ある条件を満たすデータの一括検索を容易に行なう機能を持たせることができ、この時、リンクデータは文書データとは独立したものであり、各文書に付けられているキーワードから全文書共通として自動的に生成し得るので、文書数が多い場合であっても容易にしてハイパーテキストのリンクを生成することができる。この際、請求項4記載の発明によれば、このようなリンクデータ中にユーザ毎に設定したものも選択自在に含ませたので、ユーザ毎に適切な検索が可能となる。さらに、請求項5記載の発明によれば、リンクデータを複数種のリンク集合として持たせたので、ユーザ毎、グループ毎といった単位での適切な検索を可能とすることができ、特に、このようなリンク集合を階層化させたので、リンク集合の扱いを各層単位で統一的なものとして容易化できる。また、請求項6記載の発明によれば、検索結果に応じたユーザ指示によりリンクデータ中の関連度情報を学習修正自在としたので、ユーザやグループなどの仕事や興味の変化などに対応した適切な検索を可能とすることができ。

【0046】さらに、請求項1記載の発明による場合も請求項3記載の発明による場合も、分類ツリーデータ、リンクデータが何れも文書に対するリンクを持っているので、文書検索を短時間で進めることができる。

【0047】さらに、請求項7記載の発明によれば、ユーザ登録による統制分類ツリーとは別個に、フリー分類ツリー概念を導入し、文書登録時にキーワード抽出を利用して文書を自動的にフリー分類ツリーに登録分類して、統制分類ツリーと併用自在としたので、文書の分類管理形態が細かなものとなり、これらの分類ツリーの併用により検索時の絞り込みを容易とし検索処理を短時間で済ませることができる。

【図面の簡単な説明】

【図1】本発明の第一の実施例の分類ツリー概念を示す模式図である。

【図2】全体構成を示すブロック図である。

【図3】文書データの構造を示す模式図である。

【図4】分類ツリーの文書エントリを示す模式図である。

【図5】内容一覧検索処理を示す検索画面の模式図である。

【図6】リンクデータの構造を示す模式図である。

【図7】リンク構造を示す模式図である。

【図8】検索処理の概要を示すフローチャートである。

【図9】文書表示処理の概要を示すフローチャートである。

【図10】本発明の第二の実施例のリンク構造を示す模式図である。

【図11】リンクツリー構造を示す模式図である。

【図12】本発明の第三の実施例の全体構成を示すブロック図である。

【図13】フリー分類ツリーの末端ノードの構造を示す模式図である。

【図14】フリー分類ツリーの文書エントリを示す模式*

*図である。

【図15】分類テーブルの構造を示す模式図である。

【符号の説明】

6 分類ツリーデータ記憶手段

7 リンクデータ記憶手段

8 検索処理手段

10 文書データ

11 ノード

12, 13 分類ツリー

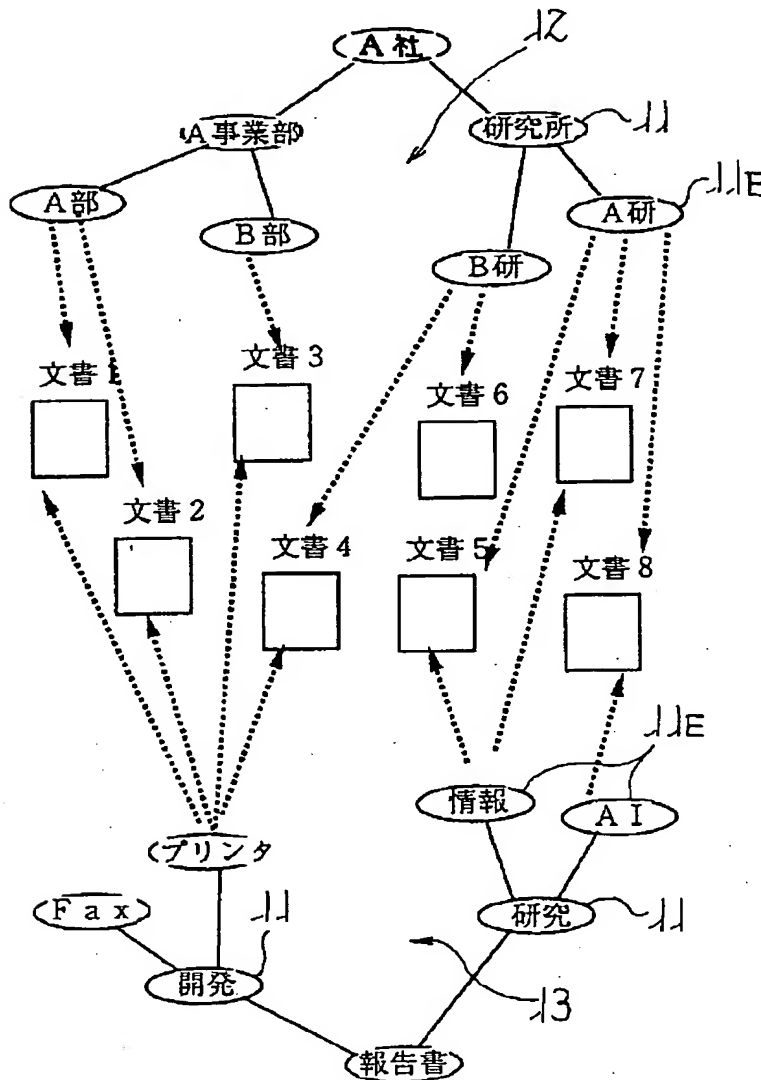
10 15 対話的処理手段

16 リンクデータ

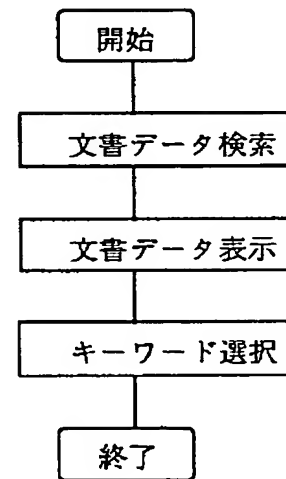
17 分類テーブル

18 文書登録手段

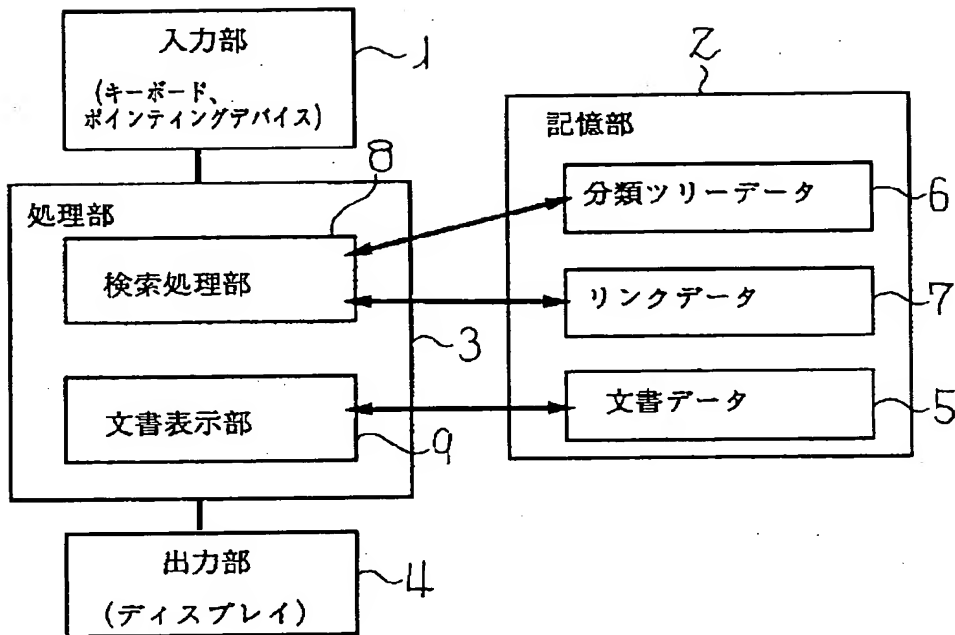
【図1】



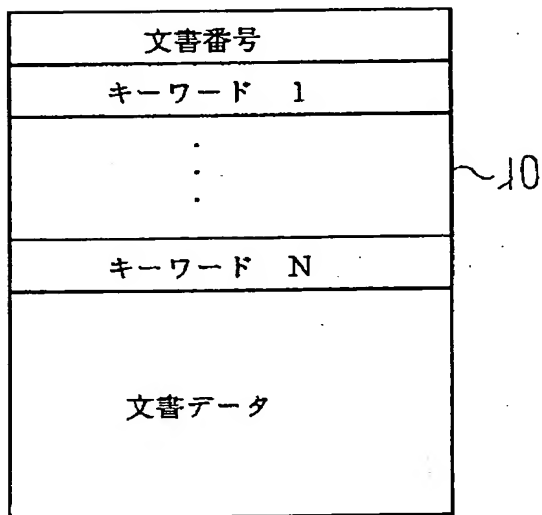
【図9】



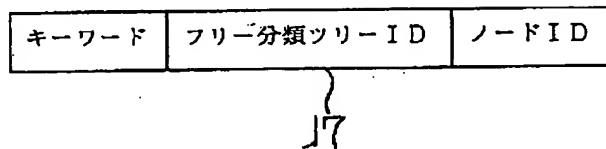
【図2】



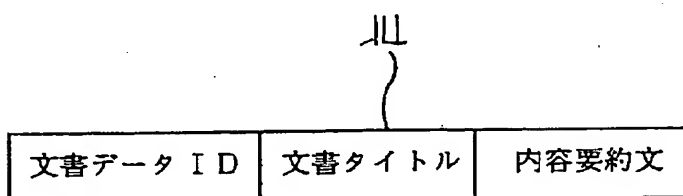
【図3】



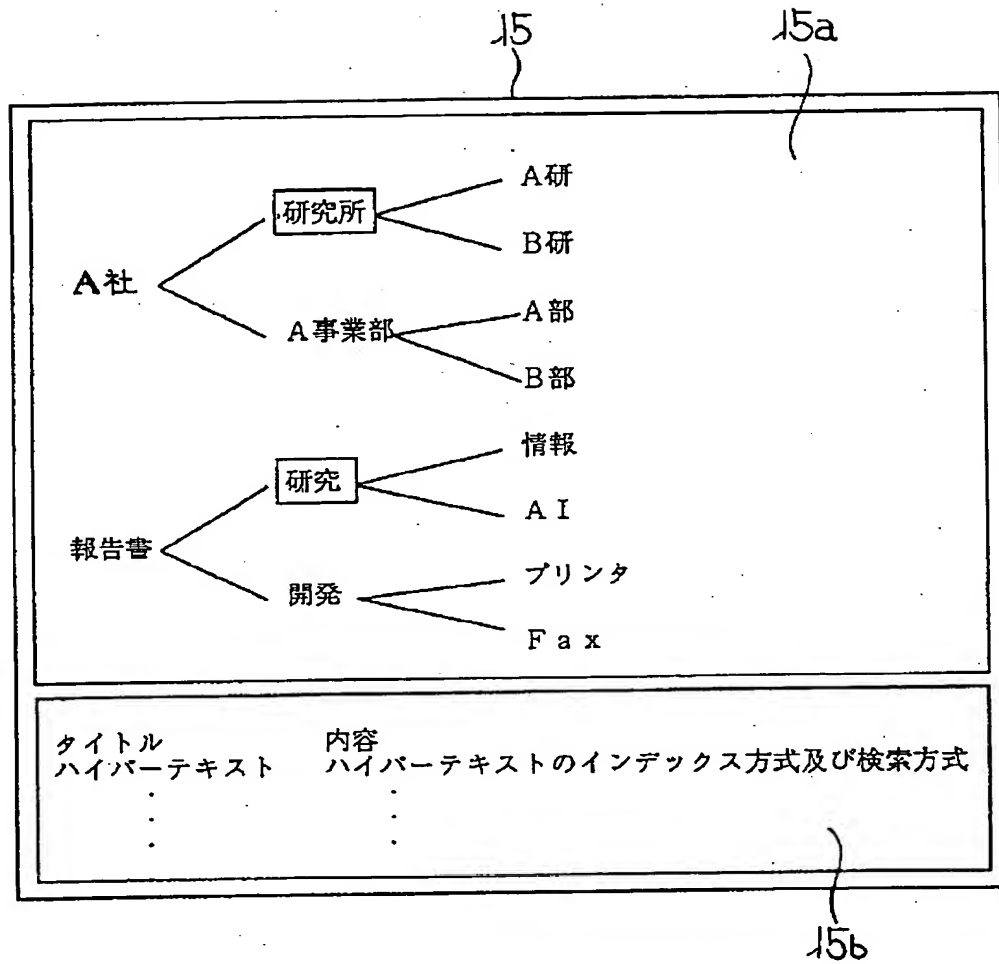
【図15】



【図4】



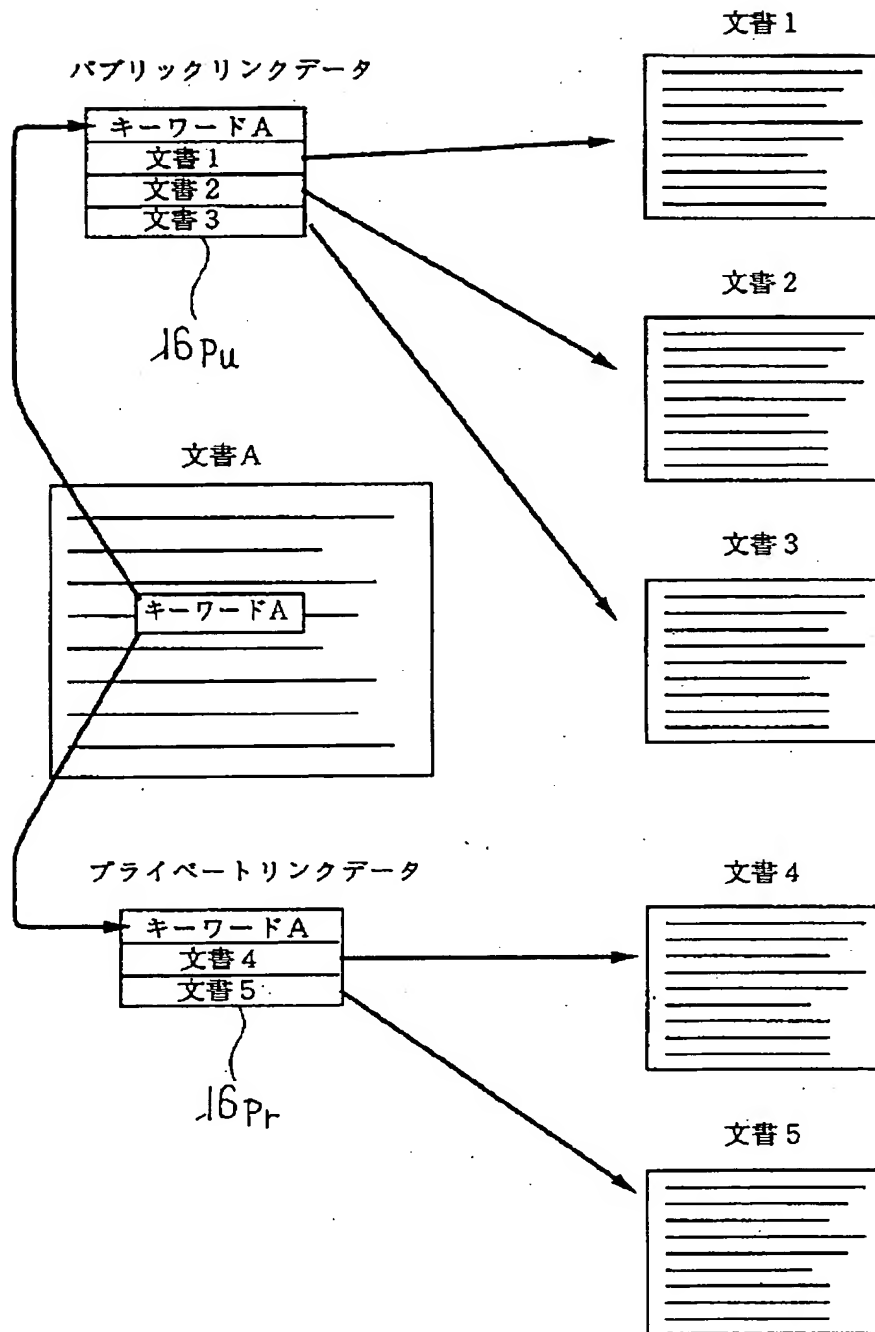
【図5】



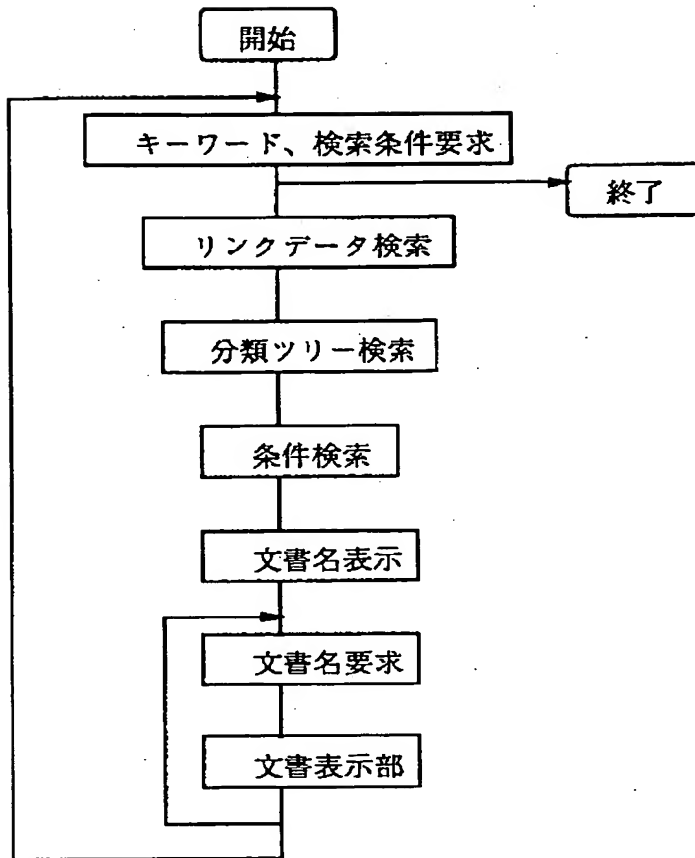
【図6】

キーワード	
文書番号 1	関連度
文書名	
⋮	
文書番号 N	関連度
文書名	

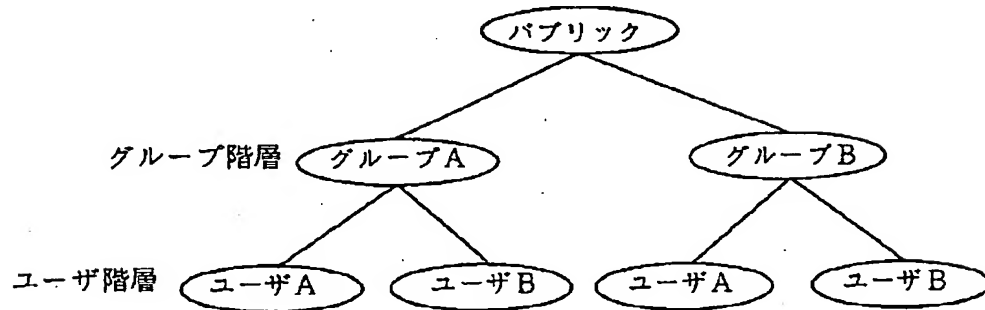
【図7】



【図8】



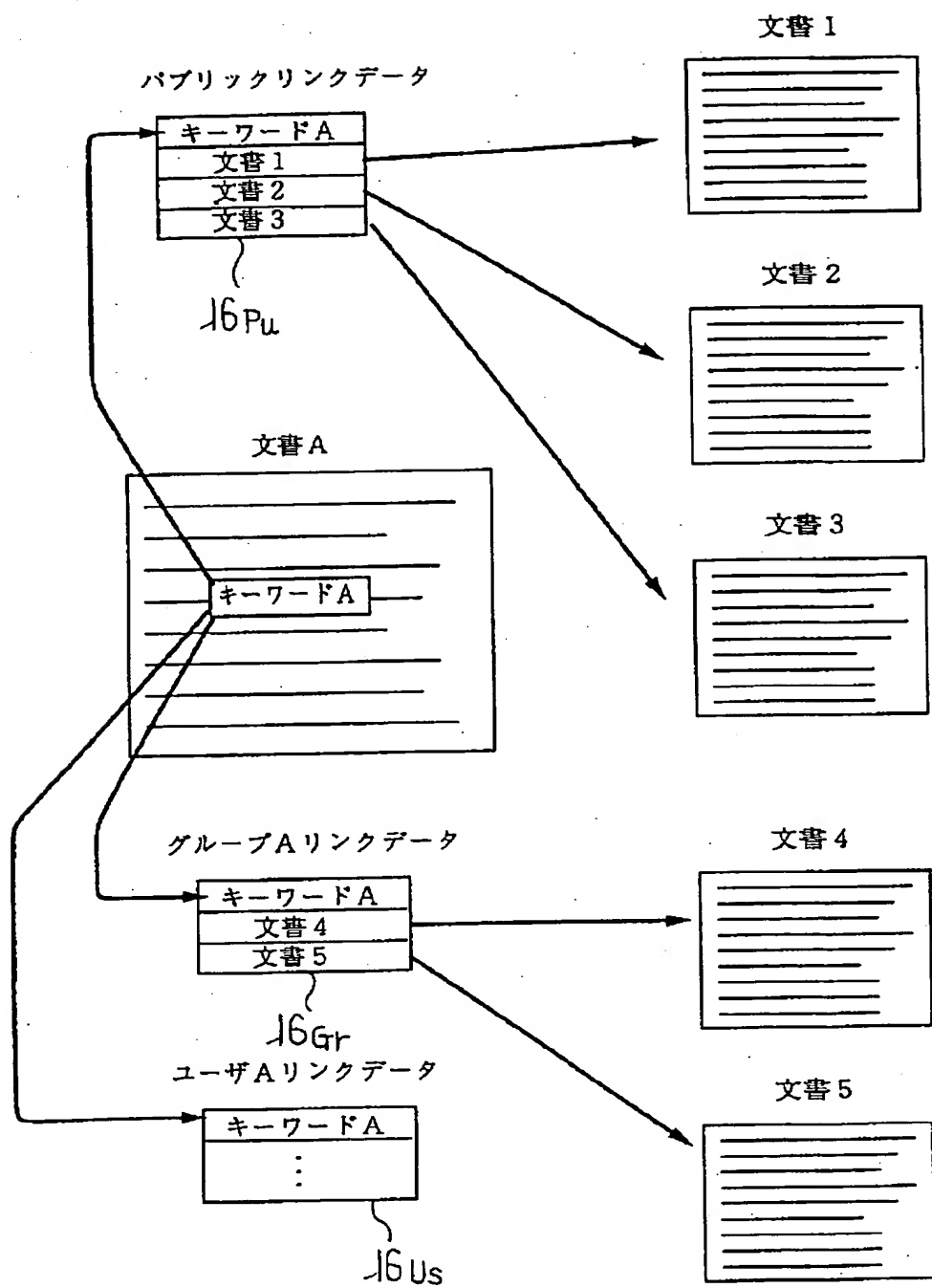
【図11】



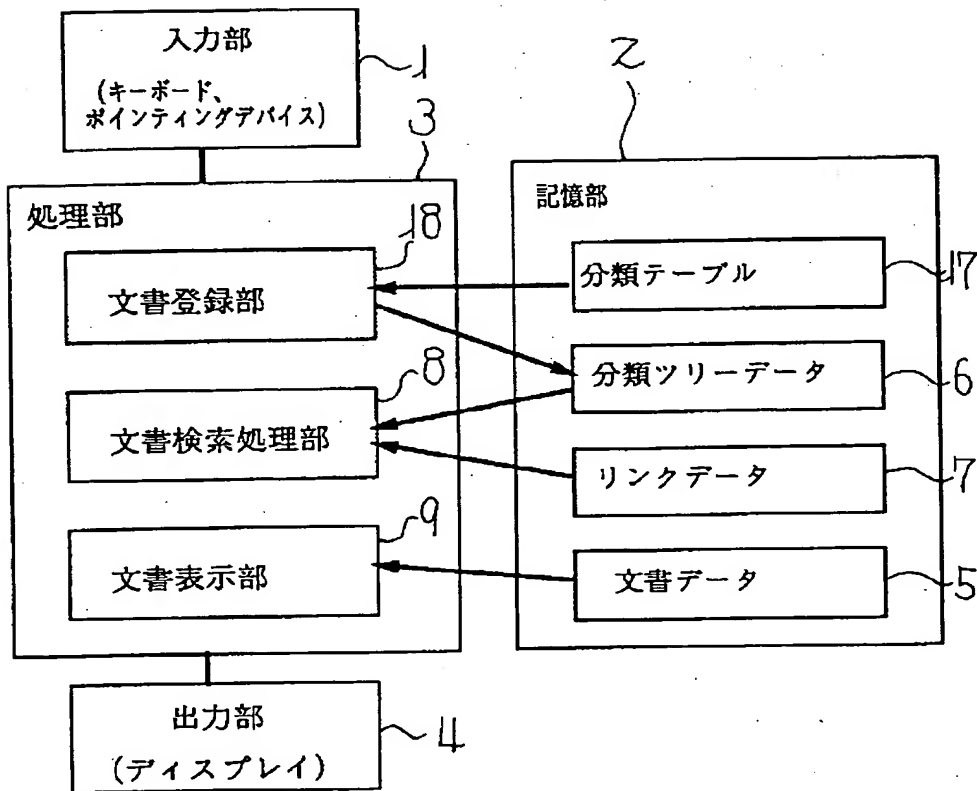
【図13】

ノード名	下位ノードリンク	文書テーブルポインタ	キーワード1	キーワード2	...
------	----------	------------	--------	--------	-----

【図10】



【図12】



【図14】

